

ECE 517: MACHINE LEARNING

ASSIGNMENT 2.2: MAXIMUM MARGIN MACHINE

DAVID KIRBY – 101652098 – DAVIDKIRBY@UNM.EDU

FALL 2021



Write a complete derivation of the proof that the maximization of margin d is equivalent to minimize the norm of the parameter vector in the maximum margin machine presented in the last set of slides of this module.

The equation for the hyperplane bisecting our data is

$$\mathbf{w}^T \mathbf{x}_0 + b = 0 \quad (1)$$

where \mathbf{w} is the norm of the parameter vector. From there, we are able to determine two support vectors on either side corresponding to our data $+1$ and -1 , as well as the margin. These equations are respectively:

$$\mathbf{w}^T \mathbf{x}_0 + b = +1$$

$$\mathbf{w}^T \mathbf{x}_0 + b = -1$$

The distance from these support vectors to the hyperplane is the margin and given by

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{x}_0 + \rho \mathbf{w} \\ d &= \|\mathbf{x}_1 - \mathbf{x}_0\| = \rho \|\mathbf{w}\| \end{aligned} \quad (2)$$

where \mathbf{x}_0 is a point on the hyperplane and \mathbf{x}_1 is a point normal to \mathbf{x}_0 and $\rho = \frac{1}{\|\mathbf{w}\|^2}$. Using equations (1) and (2), we can solve for d in terms of the norm

$$d = \rho \|\mathbf{w}\| = \frac{1}{\|\mathbf{w}\|} \quad (3)$$

From equation (3) we can see that as $\|\mathbf{w}\|$ goes to infinity, d goes to 0 and vice versa, thus proving that the maximization of margin d is equivalent to the minimization of the norm of the parameter vector.